JOURNAL OF COMBINATORIAL THEORY 10, 246–252 (1971)

Some Extremal Problems in Geometry

PAUL ERDÖS AND GEORGE PURDY

Department of Mathematics, The University of Calgary, Alberto, Canada

Communicated by the Editor-in-Chief

Received January 7, 1969

1. INTRODUCTION

Let there be given n points $X_1, ..., X_n$ in k-dimensional Euclidean space E_k. Denote by $d(X_i, X_j)$ the distance between X_i and X_j. Let $A(X_1, ..., X_n)$ be the number of distinct values of $d(X_i, X_j)$, $1 \leqslant i \leqslant j \leqslant n$. Put $f_k(n) = \min A(X_1, ..., X_n)$, where the minimum is assumed over all possible choices of $X_1, ..., X_n$. Denote by $g_k(n)$ the maximum number of solutions of $d(X_i, X_j) = a$, $1 \leqslant i < j \leqslant n$, where the maximum is to be taken over all possible choices of a and n distinct points $X_1, ..., X_n$. The estimation of $f_k(n)$ and $g_k(n)$ are difficult problems even for $k = 2$. It is known that [1, 2]:

$$c_1 n^{2/3} < f_2(n) < c_2 n / \sqrt{\log n}, \tag{1}$$

and

$$n^{1+[c_3/(\log \log n)]} < g_2(n) < c_4 n^{3/2}, \tag{2}$$

where the c's denote positive absolute constants.

It seems that in (1) the upper bound and in (2) the lower bound is close to the right order of magnitude, but we cannot even show $f_2(n) > n^{1-\epsilon}$ or $g_2(n) < n^{1+\epsilon}$.

If $k \geqslant 4$ the study of $g_k(n)$ becomes somewhat simpler [3].

A. Oppenheim asked us the question of investigating the number of triangles chosen from n points in the plane which have the same non-zero area. In this note we investigate this question and its generalizations.

2. Notations

Let $X_0, X_1, \ldots, X_n$ be n distinct points in k-dimensional space E_k, $\Delta > 0, r \geqslant 2$.

We define $g_k^{(r)}(n; X_1, \ldots, X_n; \Delta)$ $(n \geqslant r+1, k \geqslant r)$ to be the number of r-dimensional simplices of the form $X_{i_0} \cdots X_{i_r}$ having volume Δ. We let

$$g_k^{(r)}(n; X_1, \ldots, X_n) = \max_{\Delta} g_k^{(r)}(n; X_1, \ldots, X_n; \Delta)$$

and

$$g_k^{(r)}(n) = \max_{X_1, \ldots, X_n} g_k^{(r)}(n; X_1, \ldots, X_n).$$

Let X_0 be a fixed point and define

$$G_k^{(r)}(n; X_1, \ldots, X_n; \Delta)(n \geqslant r; k \geqslant r)$$

to be the number of r-dimensional simplices of the form $X_0 X_{i_1} \cdots X_{i_r}$ having volume Δ. We let

$$G_k^{(r)}(n; X_0, \ldots, X_n) = \max G_k^{(r)}(n; X_0, \ldots, X_n; \Delta)$$

and

$$G_k^{(r)}(n) = \max_{X_0, \ldots, X_n} G_k^{(r)}(n; X_0, \ldots, X_n).$$

Clearly $g_k^{(r)}(n) \leqslant nG_k^{(r)}(n-1) \leqslant nG_k^{(r)}(n)$.

We see that $g_k^{(1)}(n) = g_k(n)$ in the notation of the introduction.

We extend $f_k(n)$ to $f_k^{(r)}(n)$ and $F_k^{(r)}(n)$ in a similar way:

Let $f_k^{(r)}(n; X_1, \ldots, X_n)$ be the number of distinct volumes occurring among all the r-dimensional simplices $X_{i_0} \cdots X_{i_r}$, and let $f_k^{(r)}(n) = \min f_k^{(r)}(n; X_1, \ldots, X_n)$ where the minimum is taken over all possible choices of $X_1, \ldots, X_n$, except where $X_1, \ldots, X_n$ all lie in an $(r-1)$-dimensional subspace (not necessarily through the origin).

Similarly, if X_0 is a fixed point, let $F_k^{(r)}(n; X_0, \ldots, X_n)$ be the number of distinct volumes occurring among the r-dimensional simplices $X_0 X_{i_1} \cdots X_{i_r}$, and let $F_k^{(r)}(n) = \min F_k^{(r)}(n; X_0, \ldots, X_n)$ where the minimum is taken over $X_0, \ldots, X_n$ not lying is an $(r-1)$-dimensional subspace.

Clearly we have the following: $f_k^{(r)}(n) \leqslant nF_k^{(r)}(n-1) \leqslant nF_k^{(r)}(n)$, $f_k^{(1)}(n) = f_k(n)$ in the notation of Section 1. $g_{k-1}^{(r)}(n) \leqslant g_k^{(r)}(n)$,

$$G_{k-1}^{(r)}(n) \leqslant G_k^{(r)}(n), \ f_{k-1}^{(r)}(n) \geqslant f_k^{(r)}(n), \text{ and } F_{k-1}^{(r)}(n) \geqslant F_k^{(r)}(n) \qquad (k > r).$$

3.

Oppenheim pointed out that the generalized construction of Lenz (see, e.g., [3]) gives us lower bounds for g and G. To illustrate, we show that $G_4^{(2)}(2n) \geqslant n^2$.

Let (x_i, y_i) $(1 \leqslant i \leqslant n)$ be distinct pair of real numbers such that $x_i^2 + y_i^2 = 1$. Let $X_i = (0, 0, x_i, y_i,)$, $Y_i = (x_i, y_i, 0, 0)$, $(1 \leqslant i \leqslant n)$, $X_0 = (0, 0, 0, 0)$. Then the n^2 triangles $X_0 X_i Y_j$ are congruent and therefore have the same area.

The same method shows that $G_{2k}^{(k)}(kn) \geqslant n^k$ and $g_{2k+2}^{(k)}(kn+1) \geqslant n^{k+1}$. It seems to us that

$$g_{2k+2}^{(k)}(n) = \frac{n^{k+1}}{(k+1)^{k+1}}(1 + o(1)),$$

i.e., that Oppenheim's example is asymptotically best possible.

It also seems that

$$g_{2k+1}^{(k)}(n) = c_n{}^{k+1-\epsilon_k},$$

and we have proved this for $k = 2$, but we do not include the proof here.

THEOREM 1. *$G_2^{(2)}(n) \leqslant 4n^{3/2}$ and therefore $g_2^{(2)}(n) \leqslant 4n^{5/2}$.*

Proof. Suppose that, for some least n, $G_2^{(2)}(n) > 4n^{3/2}$. Then $n \geqslant 4$. Let

$$G_2^{(2)}(n; X_0, \ldots, X_n; \Delta) = m > 4n^{3/2}, \Delta > 0.$$

Let G be the graph whose vertices are $X_1, \ldots, X_n$ and whose edges are all the $X_i X_j$ such that the triangle $X_0 X_i X_j$ has area Δ. Then every vertex X_i of G is adjacent to at least $[4\sqrt{n}]$ other vertices, since, otherwise, removing X_i would reduce the number of triangles by at most $4\sqrt{n}$, and we would have

$$G_2^{(2)}(n-1) \geqslant 4n^{3/2} - 4\sqrt{n} > 4(n-1)^{3/2}$$

contradicting the minimal choice of n. If $1 \leqslant i \leqslant n$, therefore, there are at least $[4\sqrt{n}]$ points X_j such that the triangle $X_0 X_i X_j$ has area Δ. These points lie on two lines parallel to $X_0 X_i$. One of these linear sets of points, say S_i, contains at least $\frac{1}{2}[4\sqrt{n}]$ points. Consider the points X_i on the first $[\sqrt{n}]$ lines $S_j (1 \leqslant j \leqslant [\sqrt{n}])$. Then

$$n \geqslant \left| \bigcup_{1}^{[\sqrt{n}]} S_i \right| \geqslant \sum_{1}^{[\sqrt{n}]} | S_i | - \sum_{1 \leqslant i \leqslant j \leqslant [\sqrt{n}]} | S_i \cap S_j |$$

$$\geqslant \tfrac{1}{2}[\sqrt{n}][4\sqrt{n}] - \binom{[\sqrt{n}]}{2},$$

which is false for $n \geqslant 4$.

THEOREM 2.

$$g_2^{(2)}(n) \geqslant cn^2 \log\log n \qquad (n \geqslant n_0).$$

Proof. Let $n \geqslant n_0$, where n_0 will be chosen later. Let $a = [\sqrt{\log n}]$ and let $X_1, \ldots, X_m$ $(m < n)$ be the integral points (x, y) where $1 \leqslant x < n/a$ and $1 < y \leqslant a$. It is enough to show that $g_2^{(2)}(m; X_1, \ldots, X_m; \frac{1}{2}a!) \geqslant cn^2 \log\log n$ for $n \geqslant n_0$. Let (x_1, y_1) and (x_2, y_2) be integral points satisfying

$$1 \leqslant x_1 \leqslant \frac{1}{2}\left(\frac{n}{a} - a!\right),$$

$$x_1 < x_2 < \frac{n}{a} - a!,$$

$$1 \leqslant y_1 \leqslant \frac{a}{2},$$

$$y_1 < y_2 < a.$$

We may choose n_0 large enough so that $(n/a) - a! > (n/2a)$ for $n \geqslant n_0$. Let $d = (x_2 - x_1, y_2 - y_1)$. The $d + 1$ points (x_3, y_3) given by

$$\begin{aligned} x_3 &= x_1 + \frac{k}{d}(x_2 - x_1) + a!(y_2 - y_1)^{-1}, \\ y_3 &= y_1 + \frac{k}{d}(y_2 - y_1) \qquad (0 \leqslant k \leqslant d), \end{aligned} \tag{3}$$

are clearly among the points $X_1, \ldots, X_m$. Also

$$\begin{aligned} &1 \leqslant y_1 \leqslant y_3 \leqslant y_2 < a, \\ &1 \leqslant x_1 \leqslant x_3 \leqslant x_2 + a! < n/a. \end{aligned} \tag{4}$$

The area of the triangle (x_i, y_i) $(1 \leqslant i \leqslant 3)$ is easily seen to be $\frac{1}{2}a!$, and condition (4) ensures that no unordered triple $X_iX_jX_k$ is represented more than once in the form (x_i, y_i) $(1 \leqslant i \leqslant 3)$.

Let $0 < d < \sqrt{a}$. We choose (x_1, y_1) and (x_2, y_2) so that

$$(x_2 - x_1, y_2 - y_1) = d, \text{ i.e., } x_2 - x_1 = \mu d \text{ and } y_2 - y_1 = \nu d, (\mu, \nu) = 1,$$

i.e.,

$$1 \leqslant \mu < \frac{1}{d}\left(\frac{n}{a} - a!\right) \; 1 \leqslant \nu < \frac{a}{d}.$$

For each (μ, ν), (x_1, y_1), d, this determines (x_2, y_2). It is well known and easy to prove by elementary number theory that in a rectangle of sides t_1 and t_2 the number of points with coprime coordinates is

$$(1 + o(1)) \frac{6}{\pi^2} t_1 t_2 \qquad \text{as} \quad t_1 \to \infty \qquad \text{and} \quad t_2 \to \infty.$$

The point (x_1, y_1) can be chosen in

$$\left[\frac{a}{2}\right]\left[\frac{1}{2}\left(\frac{n}{a} - a!\right)\right] > cn \text{ ways}$$

and thus the number of choices of (x_1, y_1), (x_2, y_2) is greater than cn^2/d^2. Now on the line $(x_1, y_1)(x_2, y_2)$ there are $d + 1$ lattice points given by (3). Thus there are $d + 1$ choices for (x_3, y_3). Thus the number of triangles $(x_1, y_1)(x_2, y_2)(x_3, y_3)$, $(x_2 - x_1, y_2 - y_1) = d$ having area $\frac{1}{2}a!$ is more than $c(n^2/d)$. Summing for d we get the result.

THEOREM 3. *$G_3^{(2)}(n) \leqslant cn^{2-1/3}$ and therefore $g_3^{(2)}(n) \leqslant cn^{3-1/3}$.*

Proof. Suppose that $G_3^{(2)}(n) > cn^{2-1/3}$, for some n. Then for some $\Delta > 0$ and $X_0, \ldots, X_n$ in E_3, $G_3^{(2)}(n; X_0, \ldots, X_n, \Delta) > cn^{2-1/3}$. Let G be the graph whose vertices are $X_1, \ldots, X_n$ and whose edges are X_iX_j such that the triangle $X_0X_iX_j$ has area Δ. By a theorem of Sós, Turán, and Kövári [4], there exist Y_1, Y_2, Y_3 and $Z_1, \ldots, Z_k$ such that Y_i and Z_j are joined for $1 \leqslant i \leqslant 3$, $1 \leqslant j \leqslant k$, provided that c is sufficiently large, depending only on k. Hence three cylinders with axes X_0Y_1, X_0X_2, X_0Y_3 all contain $Z_1, \ldots, Z_k$ on their surfaces. But by elementary geometry this is impossible when k is greater than some absolute constant.

Somewhat similar methods work in higher dimensions. Using a theorem on generalized graphs proved in [5, Theorem 1] it can be shown that, e.g., $g_5^{(2)}(n) \leqslant cn^{3-\epsilon}$ for some ϵ, $0 < \epsilon < 1$, and also that $G_k^{(k)}(n) \leqslant c_k n^{k-\epsilon_k}$.

We may obtain a trivial upper bound for $f_k^{(r)}(n)$. Consider the points (x, y, z) with integer coordinates $0 \leqslant x, y, z \leqslant n^{1/3}$. There are at least n, and if XYZ are three such points, the area A of triangle XYZ is not at most

$$\frac{1}{2}\binom{3}{2} n^{2/3}.$$

Since $4A^2$ is an integer, we see that

$$f_3^{(2)}(n) \leqslant \binom{3}{2}^2 n^{4/3}.$$

The same method yields

$$f_k^{(r)}(n) \leqslant \binom{k}{r}^2 n^{2r/k}$$

(the result for $f_5^{(2)}(n)$ implies $g_5^{(2)} \geqslant cn^{11/5}$).

4.

Finally we would like to mention a few related combinatorial problems: Let there be given n points in the plane. How many quadruplats can one form so that not all the six distances should be different? It is not difficult to show that one can give n points so that there should be $cn^3 \log n$ quadruplets with not all the distances distinct, but that one cannot have $cn^{7/2}$ such quadruplets. It seems that the maximum is less than $n^{3+\epsilon}$ but we could not prove this.

A well known theorem of E Pannwitz states that in a plane set of n points of diameter 1 the maximum distance can occur at most n times and n is best possible. Similarly we can ask: Let there be given n points in the plane. How many triangles can one have which have the maximal (or minimal [non-zero]) area? Unfortunately we have only trivial results. The maximum area can occur at most cn^2 times and it can occur cn times.

Many more questions could be asked, here we state only a few of them. Let there be given n points in k-dimensional space. What is the largest set of pairwise congruent (similar) triangles[2]? What is the largest set of equilateral, (isosceles) triangles[2]?

References

1. P. Erdös, On Sets of Distances of n Points, *Amer. Math. Monthly* **53** (1946), 248–250.
2. L. Moser, On the Different Distances Determined by n points, *Amer. Math. Monthly* **59** (1952), 85–91.
3. P. Erdös, On Some Applications of Graph Theory to Geometry, *Canad. J. Math.* **19** (1967), 968–971. See also P. Erdös, On Sets of Distances of n Points in Euclidean Space, *Publ. Math. Inst. Hungar. Acad. Sci.* **5** (1960), 165–169.
4. T. Kövári, V. T. Sós, and P. Turán, On a Problem of K. Zarankiewicz, *Colloq. Math.* **3** (1954), 50–57.
5. P. Erdös, On Extremal Problems of Graphs and Generalized Graphs, *Israel J. Math.* **2** (1964), 183–190.